

Invertrons, a Class of Structurally and Functionally Related Genetic Elements That Includes Linear DNA Plasmids, Transposable Elements, and Genomes of Adeno-Type Viruses

KENJI SAKAGUCHI

Nihon Shokuhin Kako Laboratories, 3-4-1 Marunouchi, Chiyoda-ku, Tokyo 100, Japan

INTRODUCTION.....	66
STREPTOMYCES LINEAR DNA PLASMIDS pSLA1 and pSLA2.....	66
GIANT LINEAR DNA PLASMIDS FROM VARIOUS STREPTOMYCES SPECIES	67
YEAST LINEAR DNA PLASMIDS pGKL1, pGKL2, and pSKL1	68
IDENTITY OF BOTH TERMINAL SEQUENCES MAY BE CAUSED BY RECOMBINATION	68
MAIZE LINEAR DNA PLASMIDS S-1 AND S-2	68
OTHER SIMILAR PLASMIDS	68
ADENOVIRUSES AND <i>B. SUBTILIS</i> ϕ 29	69
TRANSPOSABLE ELEMENTS: MEMBERS OF THE INVERTRON GROUP.....	69
RACKET FRAME MODEL	70
RACKET FRAME MODEL FOR DNA REPLICATION, INTEGRATION, AND EXCISION OF INVERTRONS.....	70
RACKET FRAME MODEL FOR TRANSPOSITION OF TRANSPOSABLE ELEMENTS (INVERTRONS).....	71
ANOTHER TYPE OF MOBILE GENETIC ELEMENTS: RETROTRANSPOSONS.....	72
CONCLUSION	72
ACKNOWLEDGMENT	73
LITERATURE CITED	73

INTRODUCTION

Linear DNA plasmids which have identical sequences in inverted orientation at their termini were found in maize pollen cells (40, 50), *Streptomyces* species (26, 29), yeasts (20, 22, 28, 38, 53), plants (11, 48), and fungi (15, 25, 43, 55, 65, 72). They share this DNA structure with adenoviruses (62), the *Bacillus subtilis* bacteriophage ϕ 29 (12, 27, 30, 31, 68), and various transposable elements such as Ac, Ds (6, 42), and Spm (53) from maize, Tam elements from *Antirrhinum majus* (3), P-elements from *Drosophila* cells (46), and transposons from *Escherichia coli* and other bacteria (19). Their mode of DNA replication in some of the viruses (12, 27, 62) and linear plasmids (26, 68) has been studied.

Replication of these DNA elements invariably starts from terminal proteins covalently bonded to their 5' ends. For adenoviruses and phage ϕ 29, it has been proved that the DNA polymerase, terminal proteins and DNA elongation factors are encoded in the viral genome. The adenoviruses and the S-1 and S-2 linear plasmids (40, 50) from maize pollen cells are known to be integrated into and excised from chromosomal or mitochondrial DNA like mobile genetic elements. It is especially noteworthy that Jung et al. (34, 35) and Fukuhara (18) recently reported the similarity of DNA polymerase genes among these linear DNA plasmids, various animal viruses, and bacteriophages. The DNA polymerase genes in yeast and maize linear DNA plasmids, pGKL1, pGKL2, and S-1, as well as in *B. subtilis* ϕ 29, *E. coli* phages PRD 1 (34) and T4, adenovirus type 2, vaccinia virus, herpesvirus, and Epstein-Barr virus, all belong to the family B DNA polymerase gene sequences, possessing the three consensus segments, 1, 2, and 3, which are characteristic of this family of polymerases.

It is convenient to regard these structurally and function-

ally similar DNA elements as a coherent group of genetic elements; the name invertrons might be suitable to describe the group. The racket frame model postulated here relates their DNA structures to their functions. Perhaps members of this class of genetic elements are closely related.

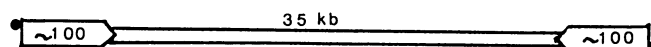
This type of mobile genetic element is distinct from the retroviruses and retro-type transposable elements like Ty1 in *Saccharomyces cerevisiae*, which have direct long terminal repeats at their ends, are homologous to the retrovirus reverse transcriptase gene, and exploit an RNA intermediate for their replication (13).

STREPTOMYCES LINEAR DNA PLASMIDS pSLA1 AND pSLA2

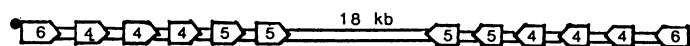
We have found two linear 17-kilobase DNA plasmids, pSLA1 and pSLA2, in *Streptomyces rochei*, which produces an antibiotic, lankacidin (26, 28, 29). These two plasmids are almost identical in DNA sequence as judged by hybridization, but pSLA1 is 500 bases shorter than pSLA2, which has been more carefully studied. At each 5' terminus of its DNA, a terminal protein is covalently attached, which perhaps serves to initiate DNA replication. The 3' ends are free, leaving them sensitive to degradation by 3' exonuclease. These properties are quite similar to those of adenovirus DNA (62) and *B. subtilis* phage ϕ 29 DNA (12). pSLA2 DNA has completely identical sequences at both termini (again similar to adeno-type viruses) for 614 base pairs (bp), followed by 11 boxes of interrupted and inverted identical repeats extending to 800 bp from both ends. In this review, this is referred to as the interrupted homologous region (Fig. 1). The racket frame model, postulated for the DNA structure of this kind of plasmid (Fig. 1), will be discussed later.

Viruses

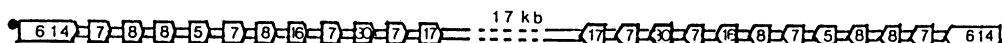
Adenoviruses
(animal cells)



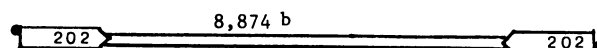
φ29
(*Bacillus subtilis*)

Linear DNA plasmids

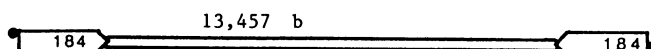
pSLA2
(*Streptomyces rochei*)



pGKL1

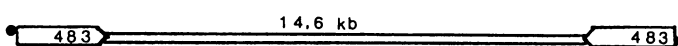


pGKL2



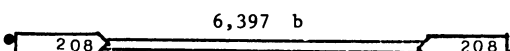
(*Kluyveromyces lactis*)

pSKL1

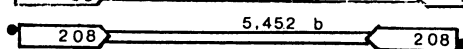


(*Saccharomyces kluyveri*)

S-1



S-2



(maize)

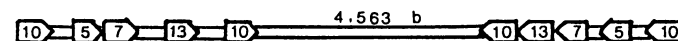
Transposable elements

transposons



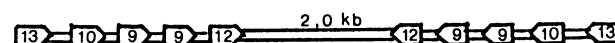
(*Escherichia coli*)

AC

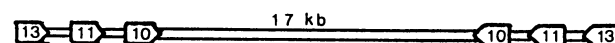


(maize)

Spm-18
(maize)

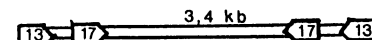


Tam 1



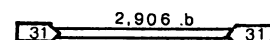
(*Antirrhinum majus*)

Le 1



(soy bean)

P-element



(*Drosophila melanogaster*)

FIG. 1. Inverttrons. The inverted terminal repeats, sometimes with the interrupted homologous regions, and terminal proteins are commonly present among various inverttrons. All the mobile genetic elements have been found only in their integrated form in the host chromosomal or mitochondrial DNA. Therefore, the possible existence of terminal proteins in a free form in the cytoplasm has not been studied. Symbol: ●, terminal proteins covalently attached to 5' ends.

GIANT LINEAR DNA PLASMIDS FROM VARIOUS *STREPTOMYCES* SPECIES

By using orthogonal field alternation gel electrophoresis, Kinashi et al. found a series of very large linear DNA plasmids (180 to 590 kilobases [kb]) from various antibiotic-producing *Streptomyces* strains; the antibiotic-producing genes are located on these plasmids (35a, 36, 37). Of these plasmids, SCP1, whose existence in *Streptomyces coelicolor* A3 (2) has been proved genetically by Wright et al. (71), was reported to be a 350-kb linear plasmid.

Kinashi et al. recently found that SCP1 DNA has inverted

repeats of more than 70 kb on both termini. They also reported that the two terminal fragments of SCP1 are attached by a protein (H. Kinashi, M. Shimaji, and T. Hanafusa, Abstr. 1990 UCLA Symposium on the Molecular Biology of *Streptomyces*. J. Cell. Biochem. Suppl.).

Hopwood and Wright have reported that SCP1 is sometimes integrated into the *S. coelicolor* chromosome, still producing the antibiotic methylenomycin A (33). These facts indicate that the giant linear plasmids from *Streptomyces* species belong to the inverttrons in their DNA structure as well as in their genetic function.

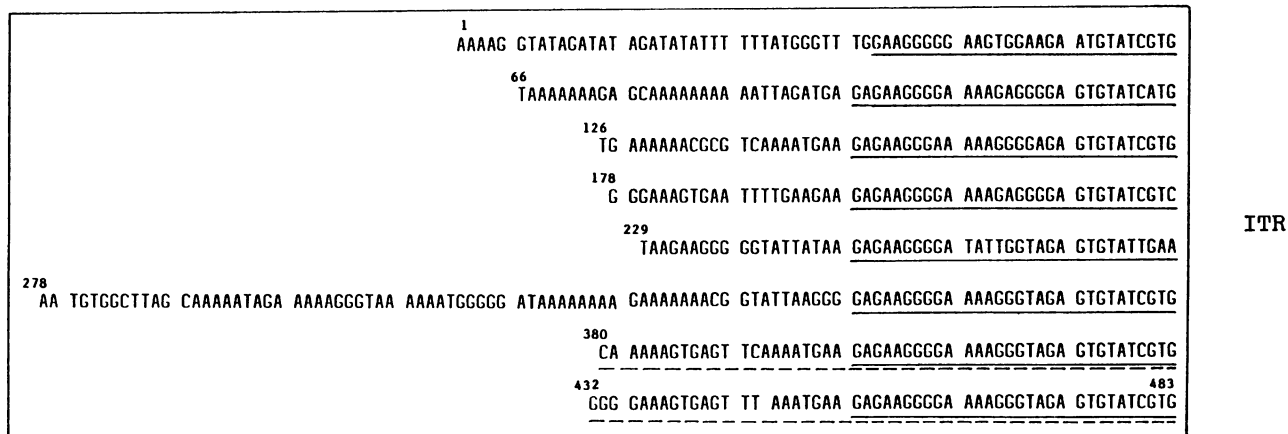


FIG. 2. Inverted terminal repeats (ITR) on the both ends of pSKL plasmid DNA (14.2 kb) from *Saccharomyces kluyveri* IFO 1685 (38). They have 483-bp identical sequences invertedly on both their ends. Within these repeats, there are eight components of the same 30-bp sequences tandemly, which suggests that the inverted terminal repeat had elongated its sequence through a slippage mechanism of DNA replication. The identity of both terminal sequences suggests the occurrence of recombination events on both termini, which requires the previous juxtaposition of both terminal sequences.

YEAST LINEAR DNA PLASMIDS pGKL1, pGKL2, AND pSKL1

Two kinds of linear DNA plasmids, pGKL1 and pGKL2 (18, 20, 22, 35, 53), 8,874 and 13,457 bp long with terminal inverted repeats of 202 and 184 bp, respectively (32, 56), were found in the yeast *Kluyveromyces lactis*. The longer plasmid, pGKL2, is responsible for the replication and maintenance of both plasmids, because the shorter one, pGKL1, is dependent on the coexistence of pGKL2 to be transformed into and propagated in *S. cerevisiae* (21, 53); on the other hand, pGKL2 can exist independently in some *S. cerevisiae* clones transformed with mixtures of the two linear plasmids. These plasmids are incompatible with the mitochondria in *S. cerevisiae* cells: they cannot replicate in the presence of mitochondria, but they can do so in *K. lactis* cells (23, 61).

The shorter pGKL1 DNA produces a protein toxin that kills *S. cerevisiae*, *Candida utilis*, *Candida tropicalis*, *Hansenula anomala*, *Kluyveromyces fragilis*, and *K. lactis* organisms that lack these two plasmids (59).

The 8,874-bp sequence of pGKL1 has been reported to have four open reading frames (ORFs) (32, 53, 57, 60); the second ORF and the leftmost ORF code for the three subunits of the toxin protein. The only ORF transcribed in the opposite direction codes for the immunity protein (63). The toxin protein acts on membrane preparations of *S. cerevisiae* in vitro and stops the synthesis of cyclic AMP (59). The rightmost ORF perhaps codes for the family B DNA polymerases which also are found in PRD and ϕ 29 phages, S-1 plasmid, adenovirus, and other animal viruses (18, 34, 35).

The longer pGKL2 has also been sequenced recently (64). It is 13,457 bp long and has 10 ORFs. ORF2 has three consensus segments with family B DNA polymerases, and ORF6 has homology with RNA polymerases (70).

Kitada and Hishinuma (38) have found a linear DNA plasmid of the same character, pSKL, from *Saccharomyces kluyveri*. It is a 14.2-kb plasmid, with identical inverted terminal sequences of 483 bp at both termini. On its 5' ends large terminal proteins are covalently bonded. It is interesting that within the inverted terminal sequences, there are eight tandemly repeated sequences of 30 bp, AAGA

GAAGGGGATATTGGTAGAGTGTATCGTG (Fig. 2). This suggests that the inverted terminal sequences of these linear plasmids arose by the process of slippage during DNA synthesis.

It is also interesting that the terminal 13-base sequence of pSKL is the same as that of pGKL2 found in *K. lactis*. The physiological function of the pSKL plasmid is not known.

pSKL has recently been sequenced by Hirai and Hishinuma (K. Hirai and F. Hishinuma, *Nucleic Acids Res.*, in press). The genome resembles that of pGKL2, having 10 ORFs with similar lengths and orientations to those in pGKL2; the DNA homology between the plasmids is 55%.

IDENTITY OF BOTH TERMINAL SEQUENCES MAY BE CAUSED BY RECOMBINATION

The complete identity of the terminal sequences of all linear DNA plasmids and adenoviruses strongly suggests that recombination events have occurred during the evolutionary process of formation of these long terminal repeats. The recombination event occurs only between the two double-stranded DNAs of same or similar sequences which have been juxtaposed closely, forming a parallel alignment of the two double strands. This property strongly supports the racket frame model discussed below.

MAIZE LINEAR DNA PLASMIDS S-1 AND S-2

Levings and co-workers (40, 50) found similar types of plasmids in sterile pollen cells of maize; these reversibly excise from and insert into the mitochondrial DNA like mobile genetic elements and adenoviruses. They have inverted identical terminal sequences of 208 bp on both their ends. At their 5' termini, terminal proteins are bonded covalently. S-2 plasmid DNA has been fully sequenced. It has 5,452 bp, with one uncharacterized ORF and one ORF which has the three consensus segments of family B DNA polymerases.

OTHER SIMILAR PLASMIDS

Similar types of linear DNA plasmids have been found in the fungi *Ascobolus immersus* (15), *Claviceps purpurea* (65),

Rhizoctonia solani (25), *Morchella conica* (43), *Pleurotus ostreatus* (72), and *Lentinula edodes* (55), and in the plants *Sorghum bicolor* (48), *Brassica campestris*, and *Brassica napus* (11). All these plasmids share the two common characteristics of the inverted terminal repeats of DNA and covalent bound terminal proteins at the 5' ends. Another characteristic property of this type of plasmid is the unusually high copy number in the host cytoplasm.

ADENOVIRUSES AND *B. SUBTILIS* $\phi 29$

Adenoviruses and *B. subtilis* $\phi 29$ have features in common with each other and with linear DNA plasmids. The DNA structure of adenoviruses and its physiological function have been well studied (5, 16, 17, 49, 72). In viral form the DNA has identical inverted terminal repeats about 100 bp long. The actual lengths of the inverted-repeat regions differ slightly among viral strains. To their 5' ends, terminal proteins of 55 kilodaltons are covalently attached. The terminal protein, DNA polymerase B family, DNA-binding protein, topoisomerase, and others are encoded in the viral genome. DNA replication starts from the 80-kilodalton precursor terminal protein; this protein attaches noncovalently to the terminus of the linear DNA, recognizing a DNA site of 10 bases very close to the end. As shown by electron microscopy, covalently bonded, newly synthesized DNA begins at the precursor terminal protein and elongates to the other terminus by displacement replacement (10).

Three other proteins of host origin, nuclear factor 1 and origin-specific binding proteins A and C, bind specifically to each of the three terminal domains, domain B (nucleotides 19 to 40 from both termini), domain A (nucleotides 1 to 18 from both termini), and domain C (nucleotides 41 to 51 from both termini), respectively. The binding of these cellular DNA-binding proteins increases the efficiency of viral DNA replication by more than 30-fold (52, 69). Thus, these site-specific binding proteins must act within the terminal inverted regions of this type of genome. The terminal proteins and the binding proteins are reported to be capable of cohesion to each other (4); this interaction adds further support to the racket frame model of function of this class of genetic elements.

When these viruses infect animal cells, their DNA can sometimes be integrated into and then excised out of the host chromosomal DNA at nonspecific sites in a manner similar to that of maize linear plasmids and transposable elements. Frolova and Zalmanzon (16) concluded that the viral DNA ends come together and insert into various sites on the cellular DNA. The sites differ among cell lines, and there is no common sequence between host and viral DNA. Together, these facts strongly support the formation of racket-frame-shaped DNA (Fig. 3); the recognition of integration sites occurs, perhaps, by protein-DNA interaction.

In some transformed cell lines the viral DNA has been completely excised and eliminated from the host genome. Still, these cells retain their oncogenic phenotype upon injection into hamsters, perhaps through previously triggered mechanisms that bring about oncogenic transformation (39).

A bacterial virus, $\phi 29$, which infects *B. subtilis* and propagates only virulently, has been studied intensively by Salas and co-workers (12, 27) and Watabe and Ito (68). The viral DNA is linear, and terminal proteins are covalently bound to the 5' termini. Hirokawa et al. (30, 31) have studied the physiological function of the terminal protein, which binds to DNA polymerase and maintains the transfectivity

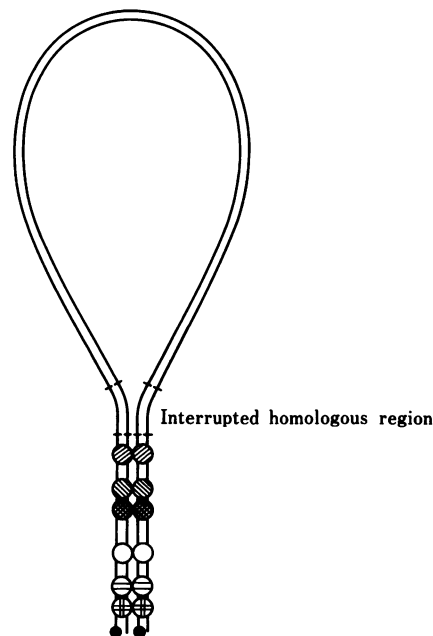


FIG. 3. Racket frame model. The inverted terminal sequences vary from 10 to 1,000 bases. However, the interrupted homologous region that is present in some cases may bind with juxtaposing proteins. Symbols: ● terminal protein which initiates DNA replication and itself has a cohesive nature; ●, ●, ●, ○, ⊖, ⊕, juxtaposing proteins which recognize and bind to the specific DNA sequences and stick together after their binding to DNA.

and replication ability of the phage genome. The DNA has 6-bp inverted terminal repeats and seven interrupted and invertedly homologous regions composed of 2- to 6-bp sequences (Fig. 1). The viral DNA codes for the terminal protein, DNA polymerase B family, DNA elongation factor, and other proteins. The terminal protein attaches to the termini of DNA by recognizing the 10 terminal bases. DNA replication initiates and elongates from protein. The terminal proteins have the property of mutual cohesion (27).

TRANSPOSABLE ELEMENTS: MEMBERS OF THE INVERTON GROUP

Many mobile genetic elements and transposons share the genetic structure and basic function with adeno-type viruses and linear DNA plasmids. These elements possess identical inverted terminal repeats, often followed by interrupted homologous regions in inverse orientation at their ends (Fig. 1).

Many transposons of *E. coli* have inverted terminal repeats of ca. 40 bp (19). The maize mobile genetic elements Ac and Ds were studied extensively by McClintock (42), and their DNA has been sequenced by Doering and Starlinger (6). They have a common terminal sequence of 11 bp and four interrupted homologous regions at their termini. Another mobile element from maize, Spm, has been studied and sequenced by Schwarz-Sommer et al. (54). It has 13-bp inverted terminal repeats with four interrupted homologous regions.

Tam 1 and certain other elements from *Antirrhinum majus* have been identified by Bonas et al. (3) as factors giving unstable changes in flower color and chalcone synthetase. Tam 1, Tam 2, and Tam 3 have 13-bp inverted terminal

sequences which are almost identical with those of Spm. It also has three interrupted homologous regions.

Two mobile genetic elements, Le 1 and Le 2, have been found by Vodkin et al. (67) in the lectin gene of soybean plants. These elements have 9-bp inverted terminal sequences followed by one interrupted homologous region composed of a 17-bp sequence.

A group of transposable elements with inverted terminal repeats has been found in a cell line of the insect *Drosophila melanogaster* (14). These P elements have been sequenced. A typical one is 2,906 bp long, with perfect inverted terminal repeats of 31 bp and four ORFs of 297, 714, 792, and 655 bp. These P elements can be inserted at various hot spots in the insect chromosomal DNA and can be excised again to restore the wild-type DNA sequence. These elements are useful genetic vectors for the molecular breeding of *Drosophila* cell lines, *Drosophila* embryos (51) and other organisms.

A transposable element of 3 kb, TU1, has been found within a histone H2B pseudogene in the sea urchin *Strongylocentrotus purpuratus* (41). This element has imperfect terminal repeats, 840 bp long, flanked by 8-bp direct repeats of the H2B gene.

Since free forms of these mobile genetic elements have never been isolated, there is no knowledge about the possible presence of terminal proteins.

Within the genomes of various transposable elements, no DNA polymerase gene has been detected. This lack may reflect the fact that these mobile genetic elements do not multiply vigorously in the cytoplasm. No genetic similarity of transposases from different transposable elements has been reported.

Adenoviruses and maize linear DNA plasmids S-1 and S-2 share with mobile genetic elements the ability to insert into and be excised from their host chromosomal or mitochondrial DNA. This characteristic and the similarities in the terminal structures could justify grouping these genetic elements within one group structurally, functionally, and evolutionally. An appropriate name for them would be invertrons. The structures of known invertrons are summarized on Fig. 1.

RACKET FRAME MODEL

We have postulated the racket frame model for replication of this kind of linear DNA genome that possesses identical sequences in inverse orientation at both ends (Fig. 3) (28, 53). The handle area has the structure of quadruple-stranded DNA composed of two double-stranded DNAs of identical sequences juxtaposed by the mediation of proteins.

The presence of site-specific binding proteins in the inverted terminal region of adenoviruses was reported by Wides et al. (69), Rosenfeld et al. (52), de Vries et al. (8), Harris and Hay (28), and Stillman et al. (58) as nuclear factor 1 and origin-specific binding proteins A and C. These authors reported that the specific binding of these proteins to the terminal area of the viral genome enhances the efficiency of DNA replication initiation more than 30-fold. They also reported that these specific binding proteins, by themselves, are cohesive. This fact, along with the cohesive nature of terminal proteins reported for adenoviruses (62) and $\phi 29$ phage (27), supports the racket frame model.

The flash-ended quadruple-stranded DNA theoretically cannot exist as pure DNA. However, such an alignment of two identical or similar double-stranded DNAs does occur in nature. One example is in the recombination of DNA. Before

the recombination occurs, juxtaposition of two similar or identical sequences of double stranded DNA must occur. A visible proof would be the *Drosophila* salivary chromosomes, in which the parallel arrangement of many strands of identical sequence, with the aid of chromatin proteins, forms the visible giant chromosome. Another example would be meiosis in eucaryotic cells, in which a pair of parental alleles juxtapose before recombination occur.

The juxtaposition of two or more identical double-stranded DNAs can occur through the mediation of juxtaposing proteins, which we have postulated (53). Juxtaposing proteins would bind to specific sites on the DNA and adhere like identical subunits of enzymes; this does occur in the case of adenovirus-specific binding proteins A, C, and factor 1. These characteristics of juxtaposing proteins allow the formation of quadruple-stranded (or more) DNA of identical sequence. Electron-microscopic studies on the replication of *B. subtilis* $\phi 29$ DNA by Hermoso and Salas (27) and of our *S. rochei* pSLA2 linear DNA plasmid showed that DNA replication always starts only from one end, never occurring from both ends before one cycle of replication has been completed. This observation is most readily comprehensible by postulating that the DNA replication machinery recognizes and binds to the quadruple-stranded DNA end where the two terminal proteins are located. The presence of a newly synthesized DNA strand could lead to structural changes in the terminal complex and perhaps prohibit further binding of the replication enzyme complex.

Levings et al. succeeded in obtaining an electron micrograph of the racket frame-like structure of the S-2 linear plasmid DNA (C. S. Levings, personal communication).

This racket frame model would also be suitable for understanding the molecular mechanism by which mobile genetic elements insert into and are excised from the host chromosomal or mitochondrial DNA (Fig. 4 and 5).

Some invertrons have shorter inverted terminal repeats. However, they always possess interrupted homologous regions, which may serve as specific binding sites for the juxtaposing proteins.

RACKET FRAME MODEL FOR DNA REPLICATION, INTEGRATION, AND EXCISION OF INVERTRONS

The racket frame model for integration and excision of invertrons has advantages over the previous straight models. In Fig. 4, a diagram of a racket frame model for replication, integration, and excision of linear DNA plasmids and adenovirus is shown. A newly synthesized proterminal protein or terminal protein encoded by the genetic element attaches to the 3' free end of the DNA, and replication starts from the protein through the specific binding of DNA-replicating machinery, a replisome to the quadruple-stranded DNA terminal area which contains two terminal proteins. DNA replication proceeds by displacement replication. Integration of many viruses and linear plasmids, which sometimes occurs into host DNA (16) at various sites, may occur through the protein-DNA recognition mechanism. This process may need the help of a ligation enzyme and a gap-filling enzyme. For the excision of the integrated DNA, the racket frame-shaped DNA structure may form through the aid of juxtaposing proteins and terminal proteins. Then cleaving enzymes or terminal proteins may release the invertron as a free form. The stabilization of integrated DNA might occur through the deletion of the opposite terminal areas from their major transcribing direction in adenoviruses (7) or alternatively by heavy mutagen-

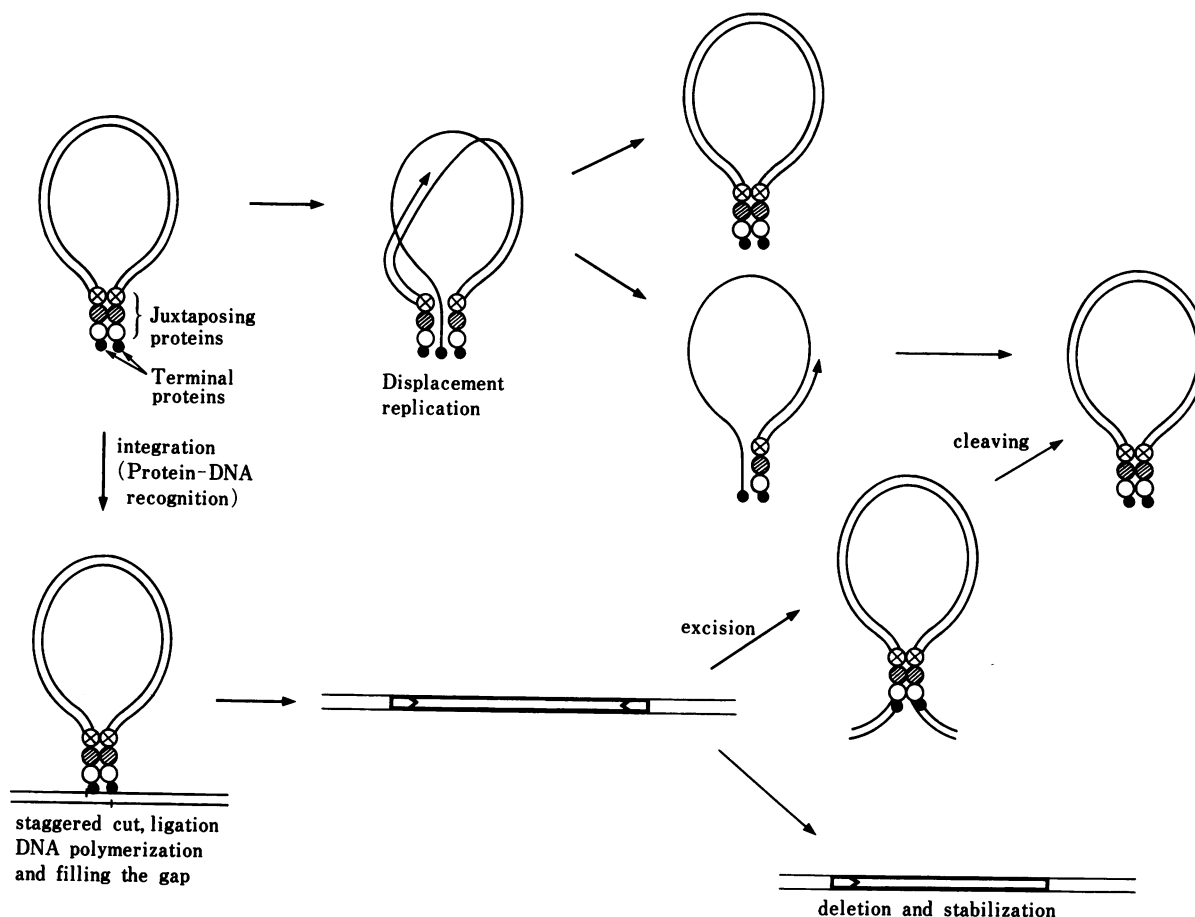


FIG. 4. Diagram of DNA replication, integration, and excision of linear DNA plasmids and adeno-type viruses based on the invertion model. See text for details.

esis in this area of the mobile genetic elements (41). Stabilization through the deletion of the right-hand terminus was also observed in many pseudogenes originating from retroviruses (1, 66) in human cells.

RACKET FRAME MODEL FOR TRANSPOSITION OF TRANSPOSABLE ELEMENTS (INVERTONS)

Figure 5 is a diagram of the transposition of transposable elements or mobile genetic elements according to the invertion model. Transposase, which is usually encoded by the transposable-element genome, binds specifically to the two identical, invertedly oriented target terminal sequences. Possibly, other juxtaposing proteins also bind to that area. With the aid of these proteins, a racket frame structure is formed and the transposase cuts one strand of each of the aligned double-stranded DNAs. DNA polymerase attaches to the free 3' ends of the cleaved DNA, and replication proceeds as far as the two racket frames of the double-stranded DNA. The resolvase, a site-specific recombination enzyme, exerts an important effect by producing one free form and another integrated form of the invertion by completion of the cointegrate type of transposition. The free-form invertion is then inserted in another target site of host DNA. In the absence of a resolvase gene in the transposon genome, the general recombination mechanism of lower efficiency can liberate the free form and mediate its integration into other sites on the host DNA (19). The existence of

a free form of the *E. coli* transposon Tn10 (19) has been reported.

In the case of the simple mechanism of insertion of transposable elements into other sites of DNA, formation of the free-form invertion by transposase cleavage of both DNA strands might play an essential role in transferring the element to other integration sites. The following points deserve special consideration.

(i) Random integration may occur at many sites. This phenomenon can be explained by the mechanism of protein-DNA interaction rather than homologous DNA recombination. The transposases or terminal proteins might be able to recognize many different integration sites.

(ii) There may be point insertion and staggered cut. Frolova and Zalmanzon (16) concluded that the right and left ends of the adenovirus DNA molecule are inserted together, as would be predicted by the racket frame model. This model can also explain more clearly the integration mechanism of transposon, especially the first step of transposition in both simple insertion and cointegration in Shapiro's model (19) and its modification by Ohtsubo et al. (19, 49), where the point insertion of transposable elements is not well understood. The invertion model may explain why the staggered cut occurs in the inserted cellular DNA. The transposase or terminal proteins bound to the two terminal parts of the mobile element would cohere and cut the acceptor DNA in a staggered manner. Their 3' ends are elongated by DNA

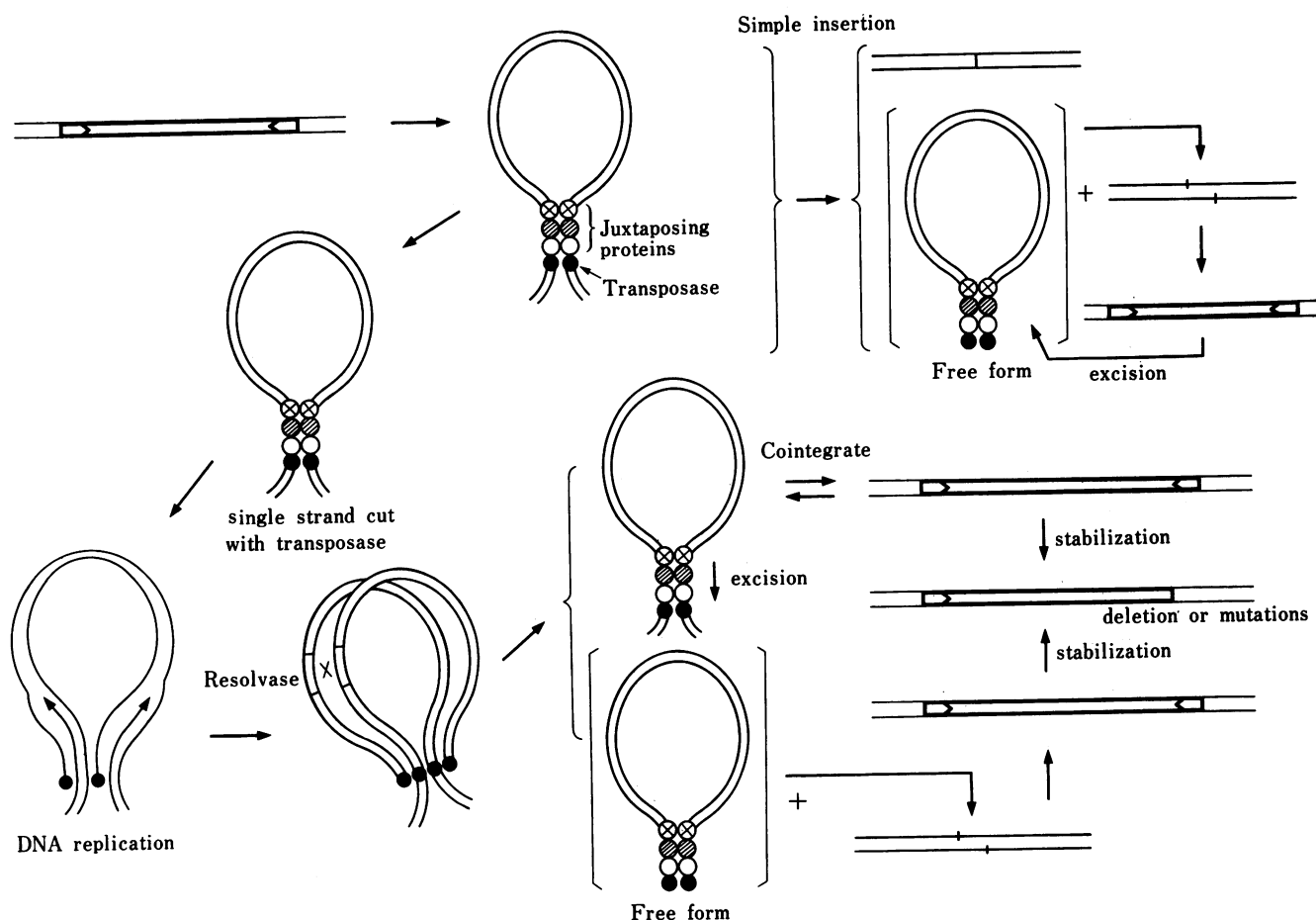


FIG. 5. Diagram of transposition of transposable elements with the invertion model. See text for details.

polymerase and ligated, leaving a short insertion (usually six or eight bases) that fills up to the gapped DNA. Of course, some transposase or terminal proteins would be able to cut the acceptor DNA such that flush ends result and no insertion segment is produced.

(iii) A transient free form may exist. For both simple insertion and cointegrate formation, the invertion model postulates a free-form invertion which exists as an intermediate just prior to integration into the target DNA.

(iv) Integrated DNA may be stabilized. The integration of the adenovirus genome into cellular DNA and its subsequent stabilization have been well studied. In a human cell line productively infected with adenovirus type 2, the intact viral genome was present in the cellular DNA (45); however, most of the stably transformed cell line exhibited deletions or inversions, or both, in the right-hand part of adenovirus type 12 DNA (7, 9), placing its major transcribing origin to the left. This also occurs with transposons. The transposon TU1, formed in multiple copies in sea urchin chromosomal DNA, had two terminal inverted sequences of around 840 bp which differed considerably. Perhaps this integrated transposon is not active.

ANOTHER TYPE OF MOBILE GENETIC ELEMENTS: RETROTRANSPOSONS

The Ty element in *S. cerevisiae* represents another type of mobile genetic element; it is 6 kb long but is flanked by direct

terminal repeats of 334 bp called δ elements. It is integrated into the yeast genome at about 30 sites, and its mRNA constitutes about 5 to 10% of the total poly(A)⁺ mRNA in a haploid yeast cell (2). Similar types of transposable elements have been found in certain eucaryotes: the *copia* elements of *Drosophila* species (44), the TED element of *Trichoplusia* (66), and retroviral proviruses (1).

These elements or proviruses share a common distinctive property: their DNA replicates via RNA through the aid of reverse transcriptase. The genes for reverse transcriptase and other genes on the retroviruses and the retrotransposons have close sequential similarity. Perhaps their ancestors are retroviruses, as reviewed by Baltimore (1). The two properties of replication via RNA intermediates and the existence of direct repeats (long terminal repeats) at both ends set retrotransposons apart from invertions as a very different class of genetic elements.

CONCLUSION

On the basis of structural and functional similarities of adeno-type viruses, linear DNA plasmids, and transposable elements, it would be possible to acknowledge the evolutionary relationship among these invertions. Moreover, it is possible to discern the traces of these genetic elements within procaryotic and eucaryotic genome structures in their integrated forms (71a). These elements have exerted important roles for the evolution of chromosomal, mitochondrial,

and other genomes through the processes of translocation, deletion, recombination, duplication, and elongation of DNA in procaryotes and eucaryotes, in parallel with the retrotransposons and retroviruses in eucaryotes and archaeobacteria.

ACKNOWLEDGMENT

I am grateful to David A. Hopwood for his reading, discussion, and correction of the manuscript.

LITERATURE CITED

- Baltimore, D. 1985. Retroviruses and retrotransposons: the role of reverse transcription in shaping the eukaryotic genome. *Cell* 40:481-482.
- Boeke, J. D., D. J. Garfinkel, C. A. Styles, and G. R. Fink. 1985. Ty elements transpose through an RNA intermediate. *Cell* 40:491-500.
- Bonas, U., H. Sommer, and H. Saedler. 1984. The 17-Kb Tam1 element of *Antirrhinum majus* induces a 3-bp duplication upon integration into the chalcone synthase gene. *EMBO J.* 3:1015-1019.
- Challberg, M. D., and T. J. Kelly. 1982. Eukaryotic DNA replication: viral and plasmid model systems. *Annu. Rev. Biochem.* 51:901-934.
- Doerfler, W. 1982. Uptake, fixation, and expression of foreign DNA in mammalian cells: the organization of integrated adenovirus DNA sequences. *Curr. Top. Microbiol. Immunol.* 101:127-194.
- Doering, H. P., and P. Starlinger. 1984. Barbara McClintock's controlling elements: now at the DNA level. *Cell* 39:253-259.
- Deuring, R., V. Winterhoff, F. Tamanai, S. Stabel, and W. Doerfler. 1981. Site of linkage between adenovirus type 12 and cell DNAs in hamster tumor line CLAC3. *Nature (London)* 293:81-84.
- de Vries, E., W. van Driel, S. J. L. van den Heuvel, and P. C. van der Vliet. 1987. Contact point analysis of the HeLa nuclear factor I: recognition site reveals symmetrical binding at one side of the DNA helix. *EMBO J.* 6:161-168.
- Eick, D., and W. Doerfler. 1982. Integrated adenovirus type 12 DNA in the transformed hamster cell line T637: sequence arrangements at the termini of viral DNA and mode of amplification. *J. Virol.* 42:317-321.
- Ellens, D. J., J. S. Sussenbach, and H. S. Jansz. 1984. Studies on the mechanism of replication of adenovirus DNA. III. Electron microscopy of replicating DNA. *Virology* 61:427-442.
- Erickson, L., W. D. Beversdorf, and K. P. Pauls. 1985. Linear mitochondrial plasmid in *Brassica* has terminal protein. *Curr. Genet.* 9:679-682.
- Escarmis, C., and M. Salas. 1981. Nucleotide sequence at the termini of the DNA of *Bacillus subtilis* phage ϕ 29. *Proc. Natl. Acad. Sci. USA* 78:1446-1450.
- Fink, G. R., J. D. Boeke, and D. J. Garfinkel. 1986. The mechanism and consequences of retrotransposition. *Trends Genet.*, p. 118-123.
- Finnegan, D. J., G. M. Rubin, M. W. Young, and D. S. Hogness. 1978. Repeated gene families in *Drosophila melanogaster*. *Cold Spring Harbor Symp. Quant. Biol.* 42:1053-1063.
- Francou, F. 1981. Isolation and characterization of a linear DNA molecule in the fungus *Ascotholus immersus*. *Mol. Gen. Genet.* 184:440-444.
- Frolova, E. I., and E. S. Zalmanzon. 1984. A study of viral genomes in cells transformed by the nononcogenic human adenovirus type 5 and highly oncogenic bovine adenovirus type 3. *Curr. Top. Microbiol. Immunol.* 111:65-89.
- Fujinaga, K., K. Yoshida, T. Yamashita, and Y. Shimizu. 1984. Organization, integration, and transcription of transforming genes of oncogenic human adenovirus type 12 and 7. *Curr. Top. Microbiol. Immunol.* 110:53-72.
- Fukuhara, H. 1987. The RF1 gene of the killer DNA of yeast may encode a DNA polymerase. *Nucleic Acids Res.* 15:10046.
- Grindley, N. D. F., and R. R. Reed. 1985. Transpositional recombination in prokaryotes. *Annu. Rev. Biochem.* 54:863-896.
- Gunge, N. 1983. Yeast DNA plasmids. *Annu. Rev. Microbiol.* 37:253-276.
- Gunge, N., K. Murata, and K. Sakaguchi. 1982. Transformation of *Saccharomyces cerevisiae* with linear DNA killer plasmids from *Kluyveromyces lactis*. *J. Bacteriol.* 151:462-464.
- Gunge, N., A. Tamaru, F. Ozawa, and K. Sakaguchi. 1981. Isolation and characterization of linear deoxyribonucleic acid plasmids from *Kluyveromyces lactis* and the plasmid-associated killer character. *J. Bacteriol.* 145:382-390.
- Gunge, N., and C. Yamane. 1984. Incompatibility of linear DNA killer plasmids pGKL1 and pGKL2 from *Kluyveromyces lactis* with mitochondrial DNA from *Saccharomyces cerevisiae*. *J. Bacteriol.* 159:533-539.
- Harris, M. P. G., and R. T. Hay. 1988. DNA sequences required for the initiation of adenovirus type 4 DNA replication in vitro. *J. Mol. Biol.* 201:57-67.
- Hashiba, T., Y. Homma, M. Hyakumachi, and I. Matsuda. 1984. Isolation of a DNA plasmid in the fungus *Rhizoctonia solani*. *J. Gen. Microbiol.* 130:2067-2070.
- Hayakawa, T., T. Tanaka, K. Sakaguchi, N. Otake, and H. Yonehara. 1979. A linear plasmid-like DNA in *Streptomyces* species producing lankacidin group antibiotics. *J. Gen. Appl. Microbiol.* 25:255-260.
- Hermoso, J. M., and M. Salas. 1980. Protein p3 is linked to the DNA of phage ϕ 29 through a phosphoester bond between serine and 5'-dAMP. *Proc. Natl. Acad. Sci. U.S.A.* 77:6425-6428.
- Hirochika, H., K. Nakamura, and K. Sakaguchi. 1984. A linear DNA plasmid from *Streptomyces rochei* with an inverted terminal repetition of 614 base pairs. *EMBO J.* 3:761-766.
- Hirochika, H., and K. Sakaguchi. 1982. Association of protein with the ends of the plasmid DNA. Analysis of linear plasmids isolated from *Streptomyces*. *Plasmid* 7:59-65.
- Hirokawa, H. 1972. Transfecting deoxyribonucleic acid of bacteriophage ϕ 29 that is protease sensitive. *Proc. Natl. Acad. Sci. USA* 69:1555-1559.
- Hirokawa, H., T. Shibata, J. Takagi, and T. Ando. 1977. Biological affinity of protease-sensitive transfecting DNAs of small size of *Bacillus* phages, p. 3-9. In A. Purtoles, R. Lopez, and M. Espinosa (ed.), *Modern trends in bacterial transformation and transfection*. North-Holland Publishing Co., Amsterdam.
- Hishinuma, F., K. Nakamura, K. Hirai, R. Nishizawa, N. Gunge, and T. Maeda. 1984. Cloning and nucleotide sequences of the linear DNA killer plasmids from yeast. *Nucleic Acids Res.* 12:7581-7597.
- Hopwood, D. A., and H. M. Wright. 1976. Interactions of the plasmid SCP1 with the chromosome of *Streptomyces coelicolor* A3(2), p. 607-619. In K. D. MacDonald (ed.), *Second International Symposium on the Genetics of Industrial Microorganisms*. Academic Press, Inc., New York.
- Jung, G., M. C. Leavitt, J. C. Hsieh, and J. Ito. 1987. Bacteriophage PRD1 DNA polymerase: evolution of DNA polymerases. *Proc. Natl. Acad. Sci. USA* 84:8287-8291.
- Jung, G., M. C. Leavitt, and J. Ito. 1987. Yeast killer plasmid pGKL1 encodes a DNA polymerase belonging to the family B DNA polymerases. *Nucleic Acids Res.* 15:9088.
- Kinashi, H. 1989. Giant linear plasmids in various methylenomycin-producing strains of *Streptomyces* species, p. 141-146. In C. L. Hershberger, S. W. Queener, and G. Hegeman (ed.), *Genetics and molecular biology of industrial microorganisms*. American Society for Microbiology, Washington, D.C.
- Kinashi, H., and M. Shimaji. 1987. Detection of giant linear plasmids in antibiotic producing strains of *Streptomyces* by the OFAGE technique. *J. Antibiot.* 40:913-916.
- Kinashi, H., M. Shimaji, and A. Sakai. 1987. Giant linear plasmids in *Streptomyces* which code for antibiotic biosynthesis genes. *Nature (London)* 328:454-456.
- Kitada, K., and F. Hishinuma. 1987. A new linear DNA plasmid isolated from the yeast *Saccharomyces kluyveri*. *Mol. Gen. Genet.* 206:377-381.
- Kuhlmann, I., and W. Doerfler. 1983. Loss of viral genomes from hamster tumor cells and nonrandom alterations in patterns of methylation of integrated adenovirus type 12 DNA. *J. Virol.*

- 47:631-636.
40. Levings, C. S., III, and R. R. Sederoff. 1983. Nucleotide sequence of the S-2 mitochondrial DNA from the S cytoplasm of maize. *Proc. Natl. Acad. Sci. USA* **80**:4055-4059.
41. Liebermann, D., B. Hoffman-Liebermann, J. Weinthal, G. Childs, R. Maxon, A. Mauron, S. N. Cohen, and L. Kedes. 1983. An unusual transposon with long terminal inverted repeats in the sea urchin *Strongylocentrotus purpuratus*. *Nature (London)* **306**:342-347.
42. McClintock, B. 1951. Chromosome organization and genic expression. *Cold Spring Harbor Symp. Quant. Biol.* **16**:13-47.
43. Meinhardt, F., and K. Esser. 1984. Linear extrachromosomal DNA in *Morchella conica*. *Curr. Genet.* **8**:15-18.
44. Miller, D. W., and L. K. Miller. 1982. A virus mutant with insertion of a copia-like transposable element. *Nature (London)* **299**:562-564.
45. Neumann, R., and W. Doerfler. 1981. Integration of adenovirus type 2 DNA at a limited number of cellular sites in productively infected cells. *J. Virol.* **37**:887-892.
46. O'Hare, K., and G. M. Rubin. 1983. Structure of P transposable elements and their sites of insertion and excision in the *Drosophila melanogaster* genome. *Cell* **34**:25-35.
47. Ohtsubo, E., M. Zenilman, H. Ohtsubo, M. McCormide, C. Machida, and Y. Machida. 1981. Mechanism of insertion and cointegration mediated by IS1 and Tn3. *Cold Spring Harbor Symp. Quant. Biol.* **45**:283-295.
48. Palmer, J. D., C. R. Shields, D. B. Cohen, and T. J. Orton. 1983. An unusual mitochondrial DNA plasmid in the genus *Brassica*. *Nature (London)* **301**:725-728.
49. Petterson, U., and R. J. Roberts. 1986. Adenovirus gene expression and replication: a historical review. *Cancer Cells* **4**:37-57.
50. Pring, D. R., C. S. Levings III, W. W. Hu, and D. H. Timothy. 1977. Unique DNA associated with mitochondria in the "S"-type cytoplasm of male-sterile maize. *Proc. Natl. Acad. Sci. USA* **74**:2904-2908.
51. Rio, D. C., F. A. Laski, and G. M. Rubin. 1986. Identification and immunochemical analysis of biologically active *Drosophila* P element transposase. *Cell* **44**:21-32.
52. Rosenfeld, P. J., E. A. O'Neill, R. J. Wides, and T. J. Kelly. 1987. Sequence-specific interactions between cellular DNA-binding protein and the adenovirus origin of DNA replication. *Mol. Cell. Biol.* **7**:875-886.
53. Sakaguchi, K., H. Hirochika, and N. Gunge. 1985. Linear plasmids with terminal inverted repeats obtained from *Streptomyces rochei* and *Kluyveromyces lactis*, p. 433-451. In D. R. Helinski and S. N. Cohen (ed.), *Plasmids in bacteria*. Plenum Publishing Corp., New York.
54. Schwarz-Sommer, Z., A. Gierl, R. B. Kloesgen, U. Wienand, P. A. Peterson, and H. Saedler. 1984. The Spm (En) transposable element controls the excision of a 2-Kb DNA insert at the wxm-8 allele of *Zea mays*. *EMBO J.* **3**:1021-1028.
55. Shishido, K., and U. Katayose. 1988. A basis in molecular breeding. *Heredity* **42**:20-24. (In Japanese.)
56. Sor, F., M. Wesolowski, and H. Fukuhara. 1983. Inverted terminal repetitions of the two linear DNAs associated with the killer character of the yeast *Kluyveromyces lactis*. *Nucleic Acids Res.* **11**:5037-5044.
57. Stark, M. J. R., and A. Boyd. 1986. The killer toxin of *Kluyveromyces lactis*: characterization of the toxin subunits and identification of the genes which encode them. *EMBO J.* **5**:1995-2002.
58. Stillman, B., J. F. X. Diffley, G. Prelich, and R. A. Guggenheim. 1986. DNA-protein interactions at the replication origins of adenovirus and SV40. *Cancer Cells* **4**:453-463.
59. Sugisaki, Y., N. Gunge, K. Sakaguchi, M. Yamasaki, and G. Tamura. 1983. *Kluyveromyces lactis* killer toxin inhibits adenylate cyclase of sensitive yeast cells. *Nature (London)* **304**:464-466.
60. Sugisaki, Y., N. Gunge, K. Sakaguchi, M. Yamasaki, and G. Tamura. 1984. Characterization of a novel killer toxin encoded by a double-stranded linear DNA plasmid of *Kluyveromyces lactis*. *Eur. J. Biochem.* **141**:1373-1375.
61. Sugisaki, Y., N. Gunge, K. Sakaguchi, M. Yamasaki, and G. Tamura. 1985. Transfer of DNA killer plasmids from *Kluyveromyces lactis* to *Kluyveromyces fragilis* and *Candida pseudotropicalis*. *J. Bacteriol.* **164**:1373-1375.
62. Tamanai, F., and B. W. Stillman. 1983. Initiation of adenovirus DNA replication in vitro requires a specific DNA sequence. *Proc. Natl. Acad. Sci. USA* **80**:6446-6450.
63. Tokunaga, M., N. Wada, and F. Hishinuma. 1987. Expression and identification of immunity determinants on linear DNA killer plasmids pGKL1 and pGKL2 in *Kluyveromyces lactis*. *Nucleic Acids Res.* **15**:1031-1046.
64. Tommasino, M., S. Ricci, and C. L. Gacottti. 1988. Genome organization of the killer plasmid pGKL2 from *Kluyveromyces lactis*. *Nucleic Acids Res.* **16**:5863-5878.
65. Tudzynski, P., and K. Esser. 1983. Extrachromosomal genetics of *Claviceps purpurea*. I. Mitochondrial DNA and mitochondrial plasmids. *Curr. Genet.* **7**:145-150.
66. Varmus, H. E. 1982. Form and function of retroviral proviruses. *Science* **216**:812-820.
67. Vodkin, L. O., P. R. Rhodes, and R. B. Goldberg. 1983. A lectin gene insertion has the structural features of a transposable element. *Cell* **34**:1023-1031.
68. Watabe, K., and J. Ito. 1983. A novel DNA polymerase induced by *Bacillus subtilis* phage ϕ 29. *Nucleic Acids Res.* **11**:8333-8342.
69. Wides, R. J., M. D. Challberg, D. R. Rawlins, and T. J. Kelly. 1987. Adenovirus origin of DNA replication: sequence requirements for replication in vitro. *Mol. Cell. Biol.* **7**:864-874.
70. Wilson, D. W., and P. A. Meacock. 1988. Extracellular gene expression in yeast: evidence for a plasmid-encoded RNA polymerase of unique structure. *Nucleic Acids Res.* **16**:8097-8112.
71. Wright, L. F., and D. A. Hopwood. 1976. Identification of the antibiotic determined by the SCP 1 plasmid of *Streptomyces coelicolor* A3(2). *J. Gen. Microbiol.* **95**:96-106.
- 71a. Yamada, T., and M. Shimaji. 1986. Peculiar feature of the organization of v-RNA genes of the *Chlorella* chloroplast DNA. *Nucleic Acids Res.* **14**:3827-3838.
72. Yui, Y., U. Katayose, and K. Shishido. 1988. Two linear plasmid-like DNA elements simultaneously maintained in *Pleurotus ostreatus*. *Biochim. Biophys. Acta* **951**:53-60.